OIPE

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/781,693A

DATE: 08/12/2002
TIME: 14:11:56

Input Set : **A:\11709-003001.TXT**
Output Set: **N:\CRF3\08122002\I781693A.raw**

```
   4 <110> APPLICANT: Chang, Tai-Jay
   6 <120> TITLE OF INVENTION: ANDROGEN RECEPTOR COMPLEX-ASSOCIATED
   7       PROTEIN
   9 <130> FILE REFERENCE: 11709-003001
  11 <140> CURRENT APPLICATION NUMBER: US 09/781,693A
C-> 12 <141> CURRENT FILING DATE: 2002-07-23
  14 <150> PRIOR APPLICATION NUMBER: US 60/262,312
  15 <151> PRIOR FILING DATE: 2001-01-17
  17 <160> NUMBER OF SEQ ID NOS: 17
  19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
  21 <210> SEQ ID NO: 1
  22 <211> LENGTH: 2580
  23 <212> TYPE: DNA
  24 <213> ORGANISM: Homo sapiens
  26 <400> SEQUENCE: 1
  27 atgtctcggg gtggctccta cccacacctg ttgtgggacg tgaggaaaag gtccctcggg        60
  28 ctggaggacc cgtcccggct gcggagtcgc tacctgggaa gaagagaatt tatccaaaga       120
  29 ttaaaacttg aagcaaccct taatgtgcat gatggttgtg ttaatacaat ctgttggaat       180
  30 gacactggag aatatatttt atctggctca gatgacacca aattagtaat tagtaatcct       240
  31 tacagcagaa aggttttgac aacaattcgt tcagggcacc gagcaaacat atttagtgca       300
  32 aagttcttac cttgtacaaa tgataaacag attgtatcct gctctggaga tggagtaata       360
  33 ttttatacca acgttgagca agatgcagaa accaacagac aatgccaatt tacgtgtcat       420
  34 tatggaacta cttatgagat tatgactgta cccaatgacc cttacacttt tctctcttgt       480
  35 ggtgaagatg gaactgttag gtggtttgat acacgcatca aaactagctg cacaaaagaa       540
  36 gattgtaaag atgatatttt aattaactgt cgacgtgctg ccacgtctgt tgctatttgc       600
  37 ccaccaatac catattacct tgctgttggt tgttctgaca gctcagtacg aatatatgat       660
  38 cggcgaatgc tgggcacaag agctacaggg aattatgcag gtcgagggac tactggaatg       720
  39 gttgcccgtt ttattccttc ccatcttaat aataagtcct gcagagtgac atctctgtgt       780
  40 tacagtgaag atggtcaaga gattctcgtt agttactctt cagattacat atatcttttt       840
  41 gacccgaaag atgatacagc acgagaactt aaaactcctt ctgcggaaga gagaagagaa       900
  42 gagttgcgac aaccaccagt taagcgtttg agacttcgtg gtgattggtc agatactgga       960
  43 cccagagcaa ggccggagag tgaacgagaa cgagatggag agcagagtcc caatgtgtca      1020
  44 ttgatgcaga gaatgtctga tatgttatca agatggtttg aagaagcaag tgaggttgca      1080
  45 caaagcaata gaggacgagg aagatctcga cccagaggtg gaacaagtca atcagatatt      1140
  46 tcaactcttc ctacggtccc atcaagtcct gatttggaag tgagtgaaac tgcaatggaa      1200
  47 gtagatactc cagctgaaca atttcttcag ccttctacat cctctacaat gtcagctcag      1260
  48 gctcattcga catcatctcc cacagaaagc cctcattcta ctcctttgct atcttctcca      1320
  49 gacagtgaac aaaggcagtc tgttgaggca tctggacacc acacacatca tcagtctgat      1380
  50 aacaataatg aaaagctgag ccccaaacca gggacaggtg aaccagtttt aagtttgcac      1440
  51 tacagcacag aaggaacaac tacaagcaca ataaaactga actttacaga tgaatggagc      1500
  52 agtatagcat caagttctag aggaattggg agccattgca atctgagggt caggaggaa       1560
  53 tctttcgtcc cacagagctc agtgcaacca ccagaaggag acagtgaaac aaaagctcct      1620
  54 gaagaatcat cagaggatgt gacaaaatat caggaaggag tatctgcaga aaacccagtt      1680
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/781,693A

DATE: 08/12/2002
TIME: 14:11:56

Input Set : A:\11709-003001.TXT
Output Set: N:\CRF3\08122002\I781693A.raw

```
55 gagaaccata tcaatataac acaatcagat aagttcacag ccaagccatt ggattccaac   1740
56 tcaggagaaa gaaatgacct caatcttgat cgctcttgtg gggttccaga agaatctgct   1800
57 tcatctgaaa aagccaagga accagaaact tcagatcaga ctagcactga gagtgctacc   1860
58 aatgaaaata acaccaatcc tgagcctcag ttccaaacag aagccactgg gccttcagct   1920
59 catgaagaaa catccaccag ggactctgct cttcaggaca cagatgacag tgatgatgac   1980
60 ccagtcctga tcccaggtgc aaggtatcga gcaggacctg gtgatagacg ctctgctgtt   2040
61 gcccgtattc aggagttctt cagacggaga aaagaaagga aagaaatgga agaattggat   2100
62 actttgaaca ttagaaggcc gctagtaaaa atggtttata aaggccatcg caactccagg   2160
63 acaatgataa aagaagccaa tttctggggt gctaactttg taatgagtgg ttctgactgt   2220
64 ggccacattt tcatctggga tcggcacact gctgagcatt tgatgcttct ggaagctgat   2280
65 aatcatgtgg taaactgcct gcagccacat ccgtttgacc caattttagc ctcatctggc   2340
66 atagattatg acataaagat ctggtcacca ttagaagagt caaggatttt taaccgaaaa   2400
67 cttgctgatg aagttataac tcgaaacgaa ctcatgctgg aagaaactag aaacaccatt   2460
68 acagttccag cctctttcat gttgaggatg ttggcttcac ttaatcatat ccgagctgac   2520
69 cggttggagg gtgacagatc agaaggctct ggtcaagaga atgaaaatga ggatgaggaa   2580
71 <210> SEQ ID NO: 2
72 <211> LENGTH: 860
73 <212> TYPE: PRT
74 <213> ORGANISM: Homo sapiens
76 <400> SEQUENCE: 2
77 Met Ser Arg Gly Gly Ser Tyr Pro His Leu Leu Trp Asp Val Arg Lys
78 1               5                  10                  15
79 Arg Ser Leu Gly Leu Glu Asp Pro Ser Arg Leu Arg Ser Arg Tyr Leu
80             20                  25                  30
81 Gly Arg Arg Glu Phe Ile Gln Arg Leu Lys Leu Glu Ala Thr Leu Asn
82         35                  40                  45
83 Val His Asp Gly Cys Val Asn Thr Ile Cys Trp Asn Asp Thr Gly Glu
84     50                  55                  60
85 Tyr Ile Leu Ser Gly Ser Asp Asp Thr Lys Leu Val Ile Ser Asn Pro
86 65                  70                  75                  80
87 Tyr Ser Arg Lys Val Leu Thr Thr Ile Arg Ser Gly His Arg Ala Asn
88             85                  90                  95
89 Ile Phe Ser Ala Lys Phe Leu Pro Cys Thr Asn Asp Lys Gln Ile Val
90             100                 105                 110
91 Ser Cys Ser Gly Asp Gly Val Ile Phe Tyr Thr Asn Val Glu Gln Asp
92             115                 120                 125
93 Ala Glu Thr Asn Arg Gln Cys Gln Phe Thr Cys His Tyr Gly Thr Thr
94     130                 135                 140
95 Tyr Glu Ile Met Thr Val Pro Asn Asp Pro Tyr Thr Phe Leu Ser Cys
96 145                 150                 155                 160
97 Gly Glu Asp Gly Thr Val Arg Trp Phe Asp Thr Arg Ile Lys Thr Ser
98             165                 170                 175
99 Cys Thr Lys Glu Asp Cys Lys Asp Asp Ile Leu Ile Asn Cys Arg Arg
100             180                 185                 190
101 Ala Ala Thr Ser Val Ala Ile Cys Pro Pro Ile Pro Tyr Tyr Leu Ala
102             195                 200                 205
103 Val Gly Cys Ser Asp Ser Ser Val Arg Ile Tyr Asp Arg Arg Met Leu
104     210                 215                 220
105 Gly Thr Arg Ala Thr Gly Asn Tyr Ala Gly Arg Gly Thr Thr Gly Met
```

RAW SEQUENCE LISTING   DATE: 08/12/2002
PATENT APPLICATION: US/09/781,693A   TIME: 14:11:56

Input Set : A:\11709-003001.TXT
Output Set: N:\CRF3\08122002\I781693A.raw

```
106 225                     230                 235                     240
107 Val Ala Arg Phe Ile Pro Ser His Leu Asn Asn Lys Ser Cys Arg Val
108                     245                 250                     255
109 Thr Ser Leu Cys Tyr Ser Glu Asp Gly Gln Glu Ile Leu Val Ser Tyr
110                     260                 265                     270
111 Ser Ser Asp Tyr Ile Tyr Leu Phe Asp Pro Lys Asp Asp Thr Ala Arg
112                     275              ·      280                 285
113 Glu Leu Lys Thr Pro Ser Ala Glu Glu Arg Arg Glu Glu Leu Arg Gln
114         290                     295                 300
115 Pro Pro Val Lys Arg Leu Arg Leu Arg Gly Asp Trp Ser Asp Thr Gly
116 305                     310                 315                     320
117 Pro Arg Ala Arg Pro Glu Ser Glu Arg Glu Arg Asp Gly Glu Gln Ser
118                     325                 330                     335
119 Pro Asn Val Ser Leu Met Gln Arg Met Ser Asp Met Leu Ser Arg Trp
120                     340                 345                     350
121 Phe Glu Glu Ala Ser Glu Val Ala Gln Ser Asn Arg Gly Arg Gly Arg
122         355                     360                 365
123 Ser Arg Pro Arg Gly Gly Thr Ser Gln Ser Asp Ile Ser Thr Leu Pro
124         370                     375                 380
125 Thr Val Pro Ser Ser Pro Asp Leu Glu Val Ser Glu Thr Ala Met Glu
126 385                     390                 395                     400
127 Val Asp Thr Pro Ala Glu Gln Phe Leu Gln Pro Ser Thr Ser Ser Thr
128                     405                 410                     415
129 Met Ser Ala Gln Ala His Ser Thr Ser Ser Pro Thr Glu Ser Pro His
130                     420                 425                     430
131 Ser Thr Pro Leu Leu Ser Ser Pro Asp Ser Glu Gln Arg Gln Ser Val
132         435                     440                 445
133 Glu Ala Ser Gly His His Thr His His Gln Ser Asp Asn Asn Asn Glu
134         450                     455                 460
135 Lys Leu Ser Pro Lys Pro Gly Thr Gly Glu Pro Val Leu Ser Leu His
136 465                     470                 475                     480
137 Tyr Ser Thr Glu Gly Thr Thr Thr Ser Thr Ile Lys Leu Asn Phe Thr
138                     485                 490                     495
139 Asp Glu Trp Ser Ser Ile Ala Ser Ser Ser Arg Gly Ile Gly Ser His
140                     500                 505                     510
141 Cys Lys Ser Glu Gly Gln Glu Glu Ser Phe Val Pro Gln Ser Ser Val
142         515                     520                 525
143 Gln Pro Pro Glu Gly Asp Ser Glu Thr Lys Ala Pro Glu Glu Ser Ser
144         530                     535                 540
145 Glu Asp Val Thr Lys Tyr Gln Glu Gly Val Ser Ala Glu Asn Pro Val
146 545                     550                 555                     560
147 Glu Asn His Ile Asn Ile Thr Gln Ser Asp Lys Phe Thr Ala Lys Pro
148                     565                 570                     575
149 Leu Asp Ser Asn Ser Gly Glu Arg Asn Asp Leu Asn Leu Asp Arg Ser
150                     580                 585                     590
151 Cys Gly Val Pro Glu Glu Ser Ala Ser Ser Glu Lys Ala Lys Glu Pro
152         595                     600                 605
153 Glu Thr Ser Asp Gln Thr Ser Thr Glu Ser Ala Thr Asn Glu Asn Asn
154         610                     615                 620
```

```
155 Thr Asn Pro Glu Pro Gln Phe Gln Thr Glu Ala Thr Gly Pro Ser Ala
156 625             630             635             Gly             640
157 His Glu Glu Thr Ser Thr Arg Asp Ser Ala Leu Gln Asp Thr Asp Asp
158             645             650             655
159 Ser Asp Asp Asp Pro Val Leu Ile Pro Gly Ala Arg Tyr Arg Ala Gly
160             660             665             670
161 Pro Gly Asp Arg Arg Ser Ala Val Ala Arg Ile Gln Glu Phe Phe Arg
162             675             680             685
163 Arg Arg Lys Glu Arg Lys Glu Met Glu Glu Leu Asp Thr Leu Asn Ile
164         690             695             700
165 Arg Arg Pro Leu Val Lys Met Val Tyr Lys Gly His Arg Asn Ser Arg
166 705             710             715             720
167 Thr Met Ile Lys Glu Ala Asn Phe Trp Gly Ala Asn Phe Val Met Ser
168             725             730             735
169 Gly Ser Asp Cys Gly His Ile Phe Ile Trp Asp Arg His Thr Ala Glu
170             740             745             750
171 His Leu Met Leu Leu Glu Ala Asp Asn His Val Val Asn Cys Leu Gln
172         755             760             765
173 Pro His Pro Phe Asp Pro Ile Leu Ala Ser Ser Gly Ile Asp Tyr Asp
174         770             775             780
175 Ile Lys Ile Trp Ser Pro Leu Glu Glu Ser Arg Ile Phe Asn Arg Lys
176 785             790             795             800
177 Leu Ala Asp Glu Val Ile Thr Arg Asn Glu Leu Met Leu Glu Glu Thr
178             805             810             815
179 Arg Asn Thr Ile Thr Val Pro Ala Ser Phe Met Leu Arg Met Leu Ala
180             820             825             830
181 Ser Leu Asn His Ile Arg Ala Asp Arg Leu Glu Gly Asp Arg Ser Glu
182         835             840             845
183 Gly Ser Gly Gln Glu Asn Glu Asn Glu Asp Glu Glu
184         850             855             860
186 <210> SEQ ID NO: 3
187 <211> LENGTH: 3016
188 <212> TYPE: DNA
189 <213> ORGANISM: Homo sapiens
191 <220> FEATURE:
192 <221> NAME/KEY: CDS
193 <222> LOCATION: (18)...(2597)
195 <400> SEQUENCE: 3
196 ccggctcagg cagagcc atg tct cgg ggt ggc tcc tac cca cac ctg ttg        50
197                     Met Ser Arg Gly Gly Ser Tyr Pro His Leu Leu
198                      1           5                   10
200 tgg gac gtg agg aaa agg tcc ctc ggg ctg gag gac ccg tcc cgg ctg       98
201 Trp Asp Val Arg Lys Arg Ser Leu Gly Leu Glu Asp Pro Ser Arg Leu
202             15              20              25
204 cgg agt cgc tac ctg gga aga aga gaa ttt atc caa aga tta aaa ctt      146
205 Arg Ser Arg Tyr Leu Gly Arg Arg Glu Phe Ile Gln Arg Leu Lys Leu
206         30              35              40
208 gaa gca acc ctt aat gtg cat gat ggt tgt gtt aat aca atc tgt tgg      194
209 Glu Ala Thr Leu Asn Val His Asp Gly Cys Val Asn Thr Ile Cys Trp
```

RAW SEQUENCE LISTING   DATE: 08/12/2002
PATENT APPLICATION: US/09/781,693A   TIME: 14:11:56

Input Set : A:\11709-003001.TXT
Output Set: N:\CRF3\08122002\I781693A.raw

```
210      45                  50                  55
212 aat gac act gga gaa tat att tta tct ggc tca gat gac acc aaa tta    242
213 Asn Asp Thr Gly Glu Tyr Ile Leu Ser Gly Ser Asp Asp Thr Lys Leu
214 60                  65                  70                  75
216 gta att agt aat cct tac agc aga aag gtt ttg aca aca att cgt tca    290
217 Val Ile Ser Asn Pro Tyr Ser Arg Lys Val Leu Thr Thr Ile Arg Ser
218                  80                  85                  90
220 ggg cac cga gca aac ata ttt agt gca aag ttc tta cct tgt aca aat    338
221 Gly His Arg Ala Asn Ile Phe Ser Ala Lys Phe Leu Pro Cys Thr Asn
222                  95                  100                 105
224 gat aaa cag att gta tcc tgc tct gga gat gga gta ata ttt tat acc    386
225 Asp Lys Gln Ile Val Ser Cys Ser Gly Asp Gly Val Ile Phe Tyr Thr
226          110                 115                 120
228 aac gtt gag caa gat gca gaa acc aac aga caa tgc caa ttt acg tgt    434
229 Asn Val Glu Gln Asp Ala Glu Thr Asn Arg Gln Cys Gln Phe Thr Cys
230          125                 130                 135
232 cat tat gga act act tat gag att atg act gta ccc aat gac cct tac    482
233 His Tyr Gly Thr Thr Tyr Glu Ile Met Thr Val Pro Asn Asp Pro Tyr
234 140                 145                 150                 155
236 act ttt ctc tct tgt ggt gaa gat gga act gtt agg tgg ttt gat aca    530
237 Thr Phe Leu Ser Cys Gly Glu Asp Gly Thr Val Arg Trp Phe Asp Thr
238                  160                 165                 170
240 cgc atc aaa act agc tgc aca aaa gaa gat tgt aaa gat gat att tta    578
241 Arg Ile Lys Thr Ser Cys Thr Lys Glu Asp Cys Lys Asp Asp Ile Leu
242          175                 180                 185
244 att aac tgt cga cgt gct gcc acg tct gtt gct att tgc cca cca ata    626
245 Ile Asn Cys Arg Arg Ala Ala Thr Ser Val Ala Ile Cys Pro Pro Ile
246          190                 195                 200
248 cca tat tac ctt gct gtt ggt tgt tct gac agc tca gta cga ata tat    674
249 Pro Tyr Tyr Leu Ala Val Gly Cys Ser Asp Ser Ser Val Arg Ile Tyr
250          205                 210                 215
252 gat cgg cga atg ctg ggc aca aga gct aca ggg aat tat gca ggt cga    722
253 Asp Arg Arg Met Leu Gly Thr Arg Ala Thr Gly Asn Tyr Ala Gly Arg
254 220                 225                 230                 235
256 ggg act act gga atg gtt gcc cgt ttt att cct tcc cat ctt aat aat    770
257 Gly Thr Thr Gly Met Val Ala Arg Phe Ile Pro Ser His Leu Asn Asn
258                  240                 245                 250
260 aag tcc tgc aga gtg aca tct ctg tgt tac agt gaa gat ggt caa gag    818
261 Lys Ser Cys Arg Val Thr Ser Leu Cys Tyr Ser Glu Asp Gly Gln Glu
262          255                 260                 265
264 att ctc gtt agt tac tct tca gat tac ata tat ctt ttt gac ccg aaa    866
265 Ile Leu Val Ser Tyr Ser Ser Asp Tyr Ile Tyr Leu Phe Asp Pro Lys
266          270                 275                 280
268 gat gat aca gca cga gaa ctt aaa act cct tct gcg gaa gag aga aga    914
269 Asp Asp Thr Ala Arg Glu Leu Lys Thr Pro Ser Ala Glu Glu Arg Arg
270          285                 290                 295
272 gaa gag ttg cga caa cca cca gtt aag cgt ttg aga ctt cgt ggt gat    962
273 Glu Glu Leu Arg Gln Pro Pro Val Lys Arg Leu Arg Leu Arg Gly Asp
274 300                 305                 310                 315
```

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/781,693A**

DATE: 08/12/2002
TIME: 14:11:57

Input Set : **A:\11709-003001.TXT**
Output Set: **N:\CRF3\08122002\I781693A.raw**

```
L:12  M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:471 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:487 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:526 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:541 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:578 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:593 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
```